

---

# **ga4gh-client Documentation**

***Release 0.0.1***

**GA4GH**

**Jan 21, 2017**



---

## Contents

---

<b>1</b>	<b>Python Client Library</b>	<b>3</b>
1.1	Types . . . . .	3
1.2	Client API . . . . .	3
<b>2</b>	<b>Release Notes</b>	<b>11</b>
2.1	0.0.3 . . . . .	11
2.2	0.0.2 . . . . .	11
2.3	0.0.1 . . . . .	11
	<b>Python Module Index</b>	<b>13</b>





# Global Alliance for Genomics & Health

`ga4gh-client` is a package that contains the client functionality of the `ga4gh` package.

See the documentation for the `ga4gh` package on its [readthedocs](#) page.

Contents:



# CHAPTER 1

---

## Python Client Library

---

This is the GA4GH client API library. This is a convenient wrapper for the low-level HTTP GA4GH API, and abstracts away network centric details such as paging. The methods and types used by the client library are defined by the [GA4GH schema](#).

**Warning:** This client API should be considered early alpha quality, and may change in arbitrary ways.

---

### Todo

A full description of this API and links to a tutorial on how to use it, as well as a quickstart showing the basic usage.

---

## 1.1 Types

---

### Todo

Add links to the upstream documentation for the GA4GH types.

---

## 1.2 Client API

---

### Todo

Add overview documentation for the client API.

---

Client classes for the GA4GH reference implementation.

---

**class** `ga4gh.client.client.AbstractClient` (*log\_level=0*)

The abstract superclass of GA4GH Client objects.

**get\_biosample** (*biosample\_id*)

Perform a get request for the given Biosample.

**Parameters** **biosample\_id** (*str*) – The ID of the Biosample

**Returns** The requested Biosample.

**Return type** `ga4gh.protocol.Biosample`

**get\_call\_set** (*call\_set\_id*)

Returns the CallSet with the specified ID from the server.

**Parameters** **call\_set\_id** (*str*) – The ID of the CallSet of interest.

**Returns** The CallSet of interest.

**Return type** `ga4gh.protocol.CallSet`

**get\_dataset** (*dataset\_id*)

Returns the Dataset with the specified ID from the server.

**Parameters** **dataset\_id** (*str*) – The ID of the Dataset of interest.

**Returns** The Dataset of interest.

**Return type** `ga4gh.protocol.Dataset`

**get\_expression\_level** (*expression\_level\_id*)

Returns the ExpressionLevel with the specified ID from the server.

**Parameters** **expression\_level\_id** (*str*) – The ID of the ExpressionLevel of interest.

**Returns** The ExpressionLevel of interest.

**Return type** `ga4gh.protocol.ExpressionLevel`

**get\_feature** (*feature\_id*)

Returns the feature with the specified ID from the server.

**Parameters** **feature\_id** (*str*) – The ID of the requested feature

**Returns** The requested `ga4gh.protocol.Feature` object.

**get\_feature\_set** (*feature\_set\_id*)

Returns the FeatureSet with the specified ID from the server.

**Parameters** **feature\_set\_id** (*str*) – The ID of the FeatureSet of interest.

**Returns** The FeatureSet of interest.

**Return type** `ga4gh.protocol.FeatureSet`

**get\_individual** (*individual\_id*)

Perform a get request for the given Individual.

**Parameters** **individual\_id** (*str*) – The ID of the Individual

**Returns** The requested Individual.

**Return type** `ga4gh.protocol.Individual`

**get\_page\_size** ()

Returns the suggested maximum size of pages of results returned by the server.



**get\_protocol\_bytes\_received()**

Returns the total number of protocol bytes received from the server by this client.

**Returns** The number of bytes consumed by protocol traffic read from the server during the lifetime of this client.

**Return type** `int`

**get\_read\_group(read\_group\_id)**

Returns the ReadGroup with the specified ID from the server.

**Parameters** `read_group_id` (*str*) – The ID of the ReadGroup of interest.

**Returns** The ReadGroup of interest.

**Return type** `ga4gh.protocol.ReadGroup`

**get\_read\_group\_set(read\_group\_set\_id)**

Returns the ReadGroupSet with the specified ID from the server.

**Parameters** `read_group_set_id` (*str*) – The ID of the ReadGroupSet of interest.

**Returns** The ReadGroupSet of interest.

**Return type** `ga4gh.protocol.ReadGroupSet`

**get\_reference(reference\_id)**

Returns the Reference with the specified ID from the server.

**Parameters** `reference_id` (*str*) – The ID of the Reference of interest.

**Returns** The Reference of interest.

**Return type** `ga4gh.protocol.Reference`

**get\_reference\_set(reference\_set\_id)**

Returns the ReferenceSet with the specified ID from the server.

**Parameters** `reference_set_id` (*str*) – The ID of the ReferenceSet of interest.

**Returns** The ReferenceSet of interest.

**Return type** `ga4gh.protocol.ReferenceSet`

**get\_rna\_quantification(rna\_quantification\_id)**

Returns the RnaQuantification with the specified ID from the server.

**Parameters** `rna_quantification_id` (*str*) – The ID of the RnaQuantification of interest.

**Returns** The RnaQuantification of interest.

**Return type** `ga4gh.protocol.RnaQuantification`

**get\_rna\_quantification\_set(rna\_quantification\_set\_id)**

Returns the RnaQuantificationSet with the specified ID from the server.

**Parameters** `rna_quantification_set_id` (*str*) – The ID of the RnaQuantificationSet of interest.

**Returns** The RnaQuantificationSet of interest.

**Return type** `ga4gh.protocol.RnaQuantificationSet`

**get\_variant(variant\_id)**

Returns the Variant with the specified ID from the server.

**Parameters** `variant_id` (*str*) – The ID of the Variant of interest.

**Returns** The Variant of interest.

**Return type** `ga4gh.protocol.Variant`

**get\_variant\_annotation\_set** (*variant\_annotation\_set\_id*)

Returns the VariantAnnotationSet with the specified ID from the server.

**Parameters** **variant\_annotation\_set\_id** (*str*) – The ID of the VariantAnnotationSet of interest.

**Returns** The VariantAnnotationSet of interest.

**Return type** `ga4gh.protocol.VariantAnnotationSet`

**get\_variant\_set** (*variant\_set\_id*)

Returns the VariantSet with the specified ID from the server.

**Parameters** **variant\_set\_id** (*str*) – The ID of the VariantSet of interest.

**Returns** The VariantSet of interest.

**Return type** `ga4gh.protocol.VariantSet`

**list\_reference\_bases** (*id\_*, *start=0*, *end=None*)

Returns an iterator over the bases from the server in the form of consecutive strings. This command does not conform to the patterns of the other search and get requests, and is implemented differently.

**search\_biosamples** (*dataset\_id*, *name=None*, *individual\_id=None*)

Returns an iterator over the Biosamples fulfilling the specified conditions.

**Parameters**

- **dataset\_id** (*str*) – The dataset to search within.
- **name** (*str*) – Only Biosamples matching the specified name will be returned.
- **individual\_id** (*str*) – Only Biosamples matching matching this id will be returned.

**Returns** An iterator over the `ga4gh.protocol.Biosample` objects defined by the query parameters.

**search\_call\_sets** (*variant\_set\_id*, *name=None*, *biosample\_id=None*)

Returns an iterator over the CallSets fulfilling the specified conditions from the specified VariantSet.

**Parameters**

- **variant\_set\_id** (*str*) – Find callsets belonging to the provided variant set.
- **name** (*str*) – Only CallSets matching the specified name will be returned.
- **biosample\_id** (*str*) – Only CallSets matching this id will be returned.

**Returns** An iterator over the `ga4gh.protocol.CallSet` objects defined by the query parameters.

**search\_datasets** ()

Returns an iterator over the Datasets on the server.

**Returns** An iterator over the `ga4gh.protocol.Dataset` objects on the server.

**search\_expression\_levels** (*rna\_quantification\_id=u'*, *feature\_ids=[]*, *threshold=0.0*)

Returns an iterator over the ExpressionLevel objects from the server

**Parameters**

- **feature\_ids** (*str*) – The IDs of the `ga4gh.protocol.Feature` of interest.

- **rna\_quantification\_id** (*str*) – The ID of the `ga4gh.protocol.RnaQuantification` of interest.
- **threshold** (*float*) – Minimum expression of responses to return.

**search\_feature\_sets** (*dataset\_id*)

Returns an iterator over the `FeatureSets` fulfilling the specified conditions from the specified `Dataset`.

**Parameters** **dataset\_id** (*str*) – The ID of the `ga4gh.protocol.Dataset` of interest.

**Returns** An iterator over the `ga4gh.protocol.FeatureSet` objects defined by the query parameters.

**search\_features** (*feature\_set\_id=None, parent\_id=u'', reference\_name=u'', start=0, end=0, feature\_types=[], name=u'', gene\_symbol=u''*)

Returns the result of running a `search_features` method on a request with the passed-in parameters.

#### Parameters

- **feature\_set\_id** (*str*) – ID of the feature Set being searched
- **parent\_id** (*str*) – ID (optional) of the parent feature
- **reference\_name** (*str*) – name of the reference to search (ex: “chr1”)
- **start** (*int*) – search start position on reference
- **end** (*int*) – end position on reference
- **feature\_types** – array of terms to limit search by (ex: “gene”)
- **name** (*str*) – only return features with this name
- **gene\_symbol** (*str*) – only return features on this gene

**Returns** an iterator over `Features` as returned in the `SearchFeaturesResponse` object.

**search\_genotype\_phenotype** (*phenotype\_association\_set\_id=None, feature\_ids=None, phenotype\_ids=None, evidence=None*)

Returns an iterator over the `GenotypePhenotype` associations from the server

**search\_individuals** (*dataset\_id, name=None*)

Returns an iterator over the `Individuals` fulfilling the specified conditions.

#### Parameters

- **dataset\_id** (*str*) – The dataset to search within.
- **name** (*str*) – Only `Individuals` matching the specified name will be returned.

**Returns** An iterator over the `ga4gh.protocol.Biosample` objects defined by the query parameters.

**search\_phenotype** (*phenotype\_association\_set\_id=None, phenotype\_id=None, description=None, type\_=None, age\_of\_onset=None*)

Returns an iterator over the `Phenotypes` from the server

**search\_phenotype\_association\_sets** (*dataset\_id*)

Returns an iterator over the `PhenotypeAssociationSets` on the server.

**search\_read\_group\_sets** (*dataset\_id, name=None, biosample\_id=None*)

Returns an iterator over the `ReadGroupSets` fulfilling the specified conditions from the specified `Dataset`.

#### Parameters

- **name** (*str*) – Only `ReadGroupSets` matching the specified name will be returned.

- **biosample\_id** (*str*) – Only ReadGroups matching the specified biosample will be included in the response.

**Returns** An iterator over the `ga4gh.protocol.ReadGroupSet` objects defined by the query parameters.

**Return type** `iter`

**search\_reads** (*read\_group\_ids*, *reference\_id=None*, *start=None*, *end=None*)

Returns an iterator over the Reads fulfilling the specified conditions from the specified *read\_group\_ids*.

**Parameters**

- **read\_group\_ids** (*str*) – The IDs of the `ga4gh.protocol.ReadGroup` of interest.
- **reference\_id** (*str*) – The name of the `ga4gh.protocol.Reference` we wish to return reads mapped to.
- **start** (*int*) – The start position (0-based) of this query. If a reference is specified, this defaults to 0. Genomic positions are non-negative integers less than reference length. Requests spanning the join of circular genomes are represented as two requests one on each side of the join (position 0).
- **end** (*int*) – The end position (0-based, exclusive) of this query. If a reference is specified, this defaults to the reference's length.

**Returns** An iterator over the `ga4gh.protocol.ReadAlignment` objects defined by the query parameters.

**Return type** `iter`

**search\_reference\_sets** (*accession=None*, *md5checksum=None*, *assembly\_id=None*)

Returns an iterator over the ReferenceSets fulfilling the specified conditions.

**Parameters**

- **accession** (*str*) – If not null, return the reference sets for which the *accession* matches this string (case-sensitive, exact match).
- **md5checksum** (*str*) – If not null, return the reference sets for which the *md5checksum* matches this string (case-sensitive, exact match). See `ga4gh.protocol.ReferenceSet.md5checksum` for details.
- **assembly\_id** (*str*) – If not null, return the reference sets for which the *assembly\_id* matches this string (case-sensitive, exact match).

**Returns** An iterator over the `ga4gh.protocol.ReferenceSet` objects defined by the query parameters.

**search\_references** (*reference\_set\_id*, *accession=None*, *md5checksum=None*)

Returns an iterator over the References fulfilling the specified conditions from the specified Dataset.

**Parameters**

- **reference\_set\_id** (*str*) – The ReferenceSet to search.
- **accession** (*str*) – If not None, return the references for which the *accession* matches this string (case-sensitive, exact match).
- **md5checksum** (*str*) – If not None, return the references for which the *md5checksum* matches this string (case-sensitive, exact match).

**Returns** An iterator over the `ga4gh.protocol.Reference` objects defined by the query parameters.

**search\_rna\_quantification\_sets** (*dataset\_id*)

Returns an iterator over the RnaQuantificationSet objects from the server

**search\_rna\_quantifications** (*rna\_quantification\_set\_id=u''*)

Returns an iterator over the RnaQuantification objects from the server

**Parameters** *rna\_quantification\_set\_id* (*str*) – The ID of the `ga4gh.protocol.RnaQuantificationSet` of interest.

**search\_variant\_annotation\_sets** (*variant\_set\_id*)

Returns an iterator over the Annotation Sets fulfilling the specified conditions from the specified variant set.

**Parameters** *variant\_set\_id* (*str*) – The ID of the `ga4gh.protocol.VariantSet` of interest.

**Returns** An iterator over the `ga4gh.protocol.AnnotationSet` objects defined by the query parameters.

**search\_variant\_annotations** (*variant\_annotation\_set\_id*, *reference\_name=u''*, *reference\_id=u''*, *start=0*, *end=0*, *effects=[]*)

Returns an iterator over the Variant Annotations fulfilling the specified conditions from the specified VariantSet.

**Parameters**

- **variant\_annotation\_set\_id** (*str*) – The ID of the `ga4gh.protocol.VariantAnnotationSet` of interest.
- **start** (*int*) – Required. The beginning of the window (0-based, inclusive) for which overlapping variants should be returned. Genomic positions are non-negative integers less than reference length. Requests spanning the join of circular genomes are represented as two requests one on each side of the join (position 0).
- **end** (*int*) – Required. The end of the window (0-based, exclusive) for which overlapping variants should be returned.
- **reference\_name** (*str*) – The name of the `ga4gh.protocol.Reference` we wish to return variants from.

**Returns** An iterator over the `ga4gh.protocol.VariantAnnotation` objects defined by the query parameters.

**Return type** `iter`

**search\_variant\_sets** (*dataset\_id*)

Returns an iterator over the VariantSets fulfilling the specified conditions from the specified Dataset.

**Parameters** *dataset\_id* (*str*) – The ID of the `ga4gh.protocol.Dataset` of interest.

**Returns** An iterator over the `ga4gh.protocol.VariantSet` objects defined by the query parameters.

**search\_variants** (*variant\_set\_id*, *start=None*, *end=None*, *reference\_name=None*, *call\_set\_ids=None*)

Returns an iterator over the Variants fulfilling the specified conditions from the specified VariantSet.

**Parameters**

- **variant\_set\_id** (*str*) – The ID of the `ga4gh.protocol.VariantSet` of interest.
- **start** (*int*) – Required. The beginning of the window (0-based, inclusive) for which overlapping variants should be returned. Genomic positions are non-negative integers less

than reference length. Requests spanning the join of circular genomes are represented as two requests one on each side of the join (position 0).

- **end** (*int*) – Required. The end of the window (0-based, exclusive) for which overlapping variants should be returned.
- **reference\_name** (*str*) – The name of the `ga4gh.protocol.Reference` we wish to return variants from.
- **call\_set\_ids** (*list*) – Only return variant calls which belong to call sets with these IDs. If an empty array, returns variants without any call objects. If null, returns all variant calls.

**Returns** An iterator over the `ga4gh.protocol.Variant` objects defined by the query parameters.

**Return type** `iter`

**set\_page\_size** (*page\_size*)

Sets the requested maximum size of pages of results returned by the server to the specified value.

**class** `ga4gh.client.client.HttpClient` (*url\_prefix*, *logLevel=30*, *authentication\_key=None*, *id\_token=None*)

The GA4GH HTTP client. This class provides methods corresponding to the GA4GH search and object GET methods.

---

## Todo

Add a better description of the role of this class and include links to the high-level API documentation.

---

## Parameters

- **urlPrefix** (*str*) – The base URL of the GA4GH server we wish to communicate with. This should include the ‘http’ or ‘https’ prefix.
- **logLevel** (*int*) – The amount of debugging information to log using the `logging` module. This is `logging.WARNING` by default.
- **authentication\_key** (*str*) – The authentication key provided by the server after logging in.
- **id\_token** (*str*) – The Auth0 `id_token` key provided by the server after logging in.

## CHAPTER 2

---

### Release Notes

---

#### **2.1 0.0.3**

Add hierarchical namespace

#### **2.2 0.0.2**

Release using new ga4gh packages, ga4gh-common and ga4gh-schemas

#### **2.3 0.0.1**

Initial release





**g**

`ga4gh.client.client`, 3

## A

AbstractClient (class in ga4gh.client.client), 3

## G

ga4gh.client.client (module), 3

get\_biosample() (ga4gh.client.client.AbstractClient method), 4

get\_call\_set() (ga4gh.client.client.AbstractClient method), 4

get\_dataset() (ga4gh.client.client.AbstractClient method), 4

get\_expression\_level() (ga4gh.client.client.AbstractClient method), 4

get\_feature() (ga4gh.client.client.AbstractClient method), 4

get\_feature\_set() (ga4gh.client.client.AbstractClient method), 4

get\_individual() (ga4gh.client.client.AbstractClient method), 4

get\_page\_size() (ga4gh.client.client.AbstractClient method), 4

get\_protocol\_bytes\_received() (ga4gh.client.client.AbstractClient method), 4

get\_read\_group() (ga4gh.client.client.AbstractClient method), 5

get\_read\_group\_set() (ga4gh.client.client.AbstractClient method), 5

get\_reference() (ga4gh.client.client.AbstractClient method), 5

get\_reference\_set() (ga4gh.client.client.AbstractClient method), 5

get\_rna\_quantification() (ga4gh.client.client.AbstractClient method), 5

get\_rna\_quantification\_set() (ga4gh.client.client.AbstractClient method), 5

get\_variant() (ga4gh.client.client.AbstractClient method), 5

get\_variant\_annotation\_set() (ga4gh.client.client.AbstractClient method), 6

get\_variant\_set() (ga4gh.client.client.AbstractClient method), 6

## H

HttpClient (class in ga4gh.client.client), 10

## L

list\_reference\_bases() (ga4gh.client.client.AbstractClient method), 6

## S

search\_biosamples() (ga4gh.client.client.AbstractClient method), 6

search\_call\_sets() (ga4gh.client.client.AbstractClient method), 6

search\_datasets() (ga4gh.client.client.AbstractClient method), 6

search\_expression\_levels() (ga4gh.client.client.AbstractClient method), 6

search\_feature\_sets() (ga4gh.client.client.AbstractClient method), 7

search\_features() (ga4gh.client.client.AbstractClient method), 7

search\_genotype\_phenotype() (ga4gh.client.client.AbstractClient method), 7

search\_individuals() (ga4gh.client.client.AbstractClient method), 7

search\_phenotype() (ga4gh.client.client.AbstractClient method), 7

search\_phenotype\_association\_sets() (ga4gh.client.client.AbstractClient method), 7

search\_read\_group\_sets() (ga4gh.client.client.AbstractClient method), 7

search\_reads() (ga4gh.client.client.AbstractClient method), 8

search\_reference\_sets() (ga4gh.client.client.AbstractClient method), 8

search\_references() (ga4gh.client.client.AbstractClient method), 8

`search_rna_quantification_sets()`  
    (ga4gh.client.client.AbstractClient method), 8

`search_rna_quantifications()`  
    (ga4gh.client.client.AbstractClient method), 9

`search_variant_annotation_sets()`  
    (ga4gh.client.client.AbstractClient method), 9

`search_variant_annotations()`  
    (ga4gh.client.client.AbstractClient method), 9

`search_variant_sets()` (ga4gh.client.client.AbstractClient  
    method), 9

`search_variants()` (ga4gh.client.client.AbstractClient  
    method), 9

`set_page_size()` (ga4gh.client.client.AbstractClient  
    method), 10